DOCUMENT RESUME

ED 338 713 TM 017 597

AUTHOR Rebarber, Theodore
TITLE Accountability in Education. Better Education through Informed Legislation Series.
INSTITUTION National Conference of State Legislatures, Washington, D.C.; Vanderbilt Univ., Nashville, TN. Educational Excellence Network.
SPONS AGENCY Donner (William H.) Foundation.; John D. and Catherine T. MacArthur Foundation, Chicago, IL.; Joyce Foundation, Chicago, IL.
REPORT NO ISBN-1-55516-325-4
PUB DATE Jul 91
NOTE 36p.; Funding also provided by the Gates Foundation.
AVAILABLE FROM National Conference of State Legislatures, 444 North Capitol Street, N.W., Suite 500, Washington, DC 20001.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Accountability; *Educational Assessment; Educational Improvement; Educational Legislation; Educational Objectives; Educational Policy; Elementary Secondary Education; *Evaluation Methods; Guides; Incentives; *State Programs; *Systems Development; Thinking Skills
IDENTIFIERS Authentic Assessment; California; Kentucky; South Carolina; Vermont

ABSTRACT

Accountability is high on the education reform agenda as taxpayers, business leaders, and elected officials demand proof of results. Although several states show promising improvements in their accountability arrangements, none has in place a comprehensive system. Such a system includes three fundamental accountability elements: (1) clear and measurable goals that describe intended outcomes; (2) assessment tools that measure progress toward these goals; and (3) incentives that reward goal-achievement and ensure adjustment in case of failure. Vermont, California, Kentucky, and other states have adopted explicit educational outcome goals for their states. South Carolina will focus efforts on two outcome goals: (1) reducing dropouts by one-half; and (2) enhancing students' higher order thinking skills. This handbook profiles South Carolina, Vermont, California, and Kentucky as they experiment with more authentic measures of student ability than traditional multiple-choice tests. Assessment approaches should measure all relevant levels of the educational system. Reforms that shift authority to a particular level increase the importance of assessment at that level. Measures for individual students are important also. South Carolina and Kentucky are implementing an incentive structure based on the performance of individual schools. Any good accountability system is tailored to the education system to which it is applied. One table is included. There are 19 references. (SLD)

TM

ED338713

Accountability in Education

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

G. Loos

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM017597

National Conference of State Legislatures

2 BEST COPY AVAILABLE

The National Conference of State Legislatures serves the legislators and staffs of the nation's 50 states, its commonwealths, and territories. NCSL was created in January 1975 from the merger of three organizations that served or represented state legislatures. NCSL is a bipartisan organization with three objectives:

To improve the quality and effectiveness of state legislatures,

To foster interstate communication and cooperation,

To ensure states a strong cohesive voice in the federal system.

The Conference has offices in Denver, Colorado and Washington, D.C.

Cover design: Jack Hadley

Accountability in Education

by

Theodor Rebarber

A joint project of the National Conference of State Legislatures
and the Vanderbilt University Educational Excellence Network

National Conference of State Legislatures
William T. Pound, Executive Director

1560 Broadway, Suite 700
Denver, Colorado 80202-5140

444 North Capitol Street, N.W., Suite 500
Washington, D.C. 20001

July 1991

The series, *Better Education Through Informed Legislation,* is printed on recycled paper.

Copyright © 1991 by the National Conference of State Legislatures. All rights reserved.
ISBN 1-55516-325-4

CONTENTS

Acknowledgments	v
Executive Summary	vii
Summary	ix
Introduction	1
Setting Goals	2
Assessment	3
New Cognitive Assessments	3
Which Levels to Measure	5
Lack of Comparative Data	5
Incentives	6
Fairness	7
Market Accountability	7
Negative Effects of Incentives	8
Cost	8
South Carolina	8
Oversight Committees	9
Assessment	9
Incentive Programs	10
Exit Examinations	10
Rewards for Schools and School Professionals	10
District Sanctions	11
EIA Results To Date	11
Target 2000: South Carolina's Second Wave of Reform	12
Vermont	12
Goals	13
Student Assessment Timeline	14
Writing Assessment	14
Mathematics Assessment	14
Local Accountability	15
California	15
Model Curriculum Frameworks	16
California Assessment Program	16
CAP Revisions	16
CAP Writing Assessment	17
New Mathematics Assessment	17
New Science and History/Social Science Assessments	17
School Report Cards	18
Other Accountability Mechanisms	19
Results to Date	19
Kentucky	19
New Bodies to Monitor Reform	20
Goal Setting	20
Assessment	21
Accountability	21
Schools	21
Districts	22
The Future in Kentucky	22
Conclusion	23
References	25
About the Author	inside back cover

ACKNOWLEDGMENTS

Our faculty included scholars and practitioners Francie Alexander, John Chubb, G. Alfred Hess, Michael Kirst, Gene Maeroff, Joseph Murphy, Thomas Payzant, and Terry Peterson, all of whom contributed to this collection. Legislators Ken Nelson of Minnesota, C. Arthur Ollie of Iowa, and Kim Peery of Washington provided the link between the realms of educational theory and legislative reality.

Floretta McKenzie, managing partner of the McKenzie Group, and former superintendent of the District of Columbia Public Schools; and testing critic John J. Cannell, author of *Nationally Normed Elementary Achievement Testing In America's Public Schools: How All Fifty States Are Above the National Average*, both addressed the institute in July, 1990.

John L. Myers, NCSL education program director, oversaw this project from inception to conclusion with the assistance of NCSL education staff, Connie Koprowicz, Tim Storey, Julie Davis Bell, Terry Whitney, and Veronica O. White. Chester E. Finn, Jr. and Theodor Rebarber of the Vanderbilt University Educational Excellence Network conducted the institute in July, 1990 and designed the handbooks and other project details with the assistance of Network staff, Andrew Forsaith, Matthew Gandal, and Courtney Uhler. Mr. Rebarber wrote the *Parent Enabling Policies* and *Accountability in Education* volumes.

This project was made possible by funding from the Donner Foundation, the Gates Foundation, the Joyce Foundation, and the John D. and Catherine T. MacArthur Foundation, all of which are committed to improving the education of our nation's children.

EXECUTIVE SUMMARY

In *A Nation at Risk*, then-Secretary of Education,Terrel Bell, warned that the deplorable state of American education endangers both our standing in the global economy and our domestic peace and civic culture. He likened it to an act of war by a hostile power. Most states responded with genuine efforts to enhance the quality of public schools. "First-wave" reforms focused on measurable standards such as stricter graduation requirements, higher teacher salaries, and minimum competency tests for teachers.

At the end of the 1990s, we find little effect from these reforms. Overall, seventeen-year olds are learning no more today than 20 years ago; despite much effort and a good deal of money spent, academic performance for the majority of students remains stagnant. Though some improvements exist among minorities, a substantial gap remains between them and most students, and the number of all pupils achieving at the highest levels has actually decreased slightly.

Current levels are far below what today's students will need in the society of the twenty-first century, too low even for today's computer-driven economy. On international comparisons, American students are consistently outperformed by students in nearly all industrialized nations, especially Germany, Japan, and South Korea.

As we approach what some call the "second wave" of reform since *Nation at Risk*, new proposals that seek to address fundamental inadequacies in the current system are before many state legislatures.

With the help of his education secretary, Lamar Alexander, President Bush recently proposed *America 2000*, a multi-faceted reform plan for the nation. The strategy includes a number of ideas with which states and localities have experimented in recent years—outcome-based national assessments tagged to world-class standards, academies in each state that aim to improve the training of teachers and principals, assistance to implement choice plans, model restructured schools for each congressional district.

When considering these ideas, most of which are voluntary, legislators need to understand the underlying principles and rationales for each. Discussion of these concepts is contained in these three handbooks. They are a resource for busy elected officials, their aides, and interested citizens trying to make sense of complex—sometimes controversial—policy questions.

Better Education Through Informed Legislation, the final product of a two-year project to inform legislators about major education-reform topics, is co-sponsored by the National Conference of State Legislatures, a nonprofit, bipartisan organization that serves state legislatures; and the Vanderbilt University Educational Excellence Network, an association of more than 1700 reform-minded educators, scholars, policymakers, and journalists.

Better Education Through Informed Legislation emphasizes reform at the state level and reflects a shift in the focus of education policy making begun in the late 1970s. While opportunities for innovation and initiative at the local level remain, state legislators are central to state education policy making and most observers agree that states will lead policy reform in the 1990s.

SUMMARY

Accountability is high on the education reform agenda as taxpayers, business leaders, and elected officials demand proof of results. Although a number of states show promising improvements in their accountability arrangements, none has in place a comprehensive system. Such a system includes three fundamental accountability elements:

- Clear and measurable goals that describe intended outcomes,
- Assessment tools that measure progress toward the goals,
- Incentives that reward goal-achievement and ensure adjustments in case of failure.

Vermont, California, Kentucky, and other states have adopted explicit outcome goals for education in their states. South Carolina will focus its efforts in the next decade on two specific outcome goals: to reduce dropouts by one half and to enhance students' higher order thinking skills. The four states profiled in this handbook—Vermont, California, South Carolina, and Kentucky—are experimenting with more "authentic" measures of student ability than traditional multiple-choice tests that assess only basic skills and facts. In addition to such efforts, current multiple-choice tests can be improved significantly.

Assessment mechanisms should measure all relevant levels of the educational system, including state, local districts, and individual schools. Reforms that shift added authority to a particular level (such as site-based management) increase the importance of quality assessments at that level.

Measures for individual students are important as well, so that students, parents, and teachers can adjust their behavior based on reliable evidence in relation to well-formulated goals and benchmarks as well as to other students.

An incentive structure that encourages progress toward goals is the final element in an accountability system. South Carolina provides financial rewards for schools, principals, and teachers whose students succeed. It also provides aid and sanctions to poorly performing districts. Similarly, Kentucky has a system of positive and negative consequences based on the performance of individual schools.
Any good accountability system provides usable information to policymakers at all relevant levels. Explicit goals set a clear task, assessments measure results, and consequences foster success.

INTRODUCTION

The current climate of education reform underscores the central role of accountability in policymaking. As part of the strong nationwide effort to revitalize our schools over the past decade—amid signs that this effort may accelerate in the years ahead—citizens demonstrate greater determination to ensure that their sizable investment in education is well spent. Taxpayers want value for their money. They want results. Legislators and other elected officials feel pressure to hold schools accountable for improved and efficient performance.

Business leaders play a more prominent role in school reform efforts, not only by demanding change, but by suggesting what changes are necessary. They bring a keen interest in education's "bottom line"—how much do students know and how capable are they of putting what they know to work?

A number of respected national organizations and commentators lament the inadequate condition of assessment systems in most states. The National Governors' Association (1989, p. 80) finds that "indicators of the quality and effectiveness of American education have consistently been lacking, especially at the state and local levels." The Council of Chief State School Officers (1989, p. 67) cites "major gaps" in the information base and notes that "outcome data meeting rigorous technical standards are not presently available." Without adequate indicators even for some of the most basic educational outcomes, states find there is an enormous amount of work to be done.

The clamor for assessment forced policymakers to rethink basic questions: What are education's underlying goals? Do current means measure the degree to which we attain these goals; and, if not, how should we alter the measures? What most effectively stimulates the best work from educators and students; how can incentives be adjusted to strengthen their efforts?

It is not possible to address the issue of incentives without considering which yardsticks determine success or failure and what these yardsticks are to assess; that is, which measurable goals? Any workable accountability system addresses the essential elements of accountability:

- Clear and measurable outcome goals,
- Reliable sources of information on progress toward the goals,
- Consequences linked to performance.

Accountability in Education explores these three topics in depth and addresses several related issues, including cost, test reliability and security, and accountability's relation to other reforms. The introduction suggests a framework for thinking about accountability. Subsequent sections examine promising or innovative reform efforts in four states. No jurisdiction has a functioning system that successfully addresses all aspects of accountability; however, several have major components, and a few are moving toward a comprehensive system.

Setting Goals

The first step to designing an accountability system is to set clear and measurable goals that spell out the results or outcomes sought. The fundamental goal is a clear definition of the "well educated" individual. Policymakers must decide what they want a high school graduate to know and be able to do. How well will he understand the basic principles of science and the physical world? To what level should he know history, geography, civics, and economics? Should he acquire an appreciation for the aesthetic and creative arts, including literature, music, and the visual and performing arts? Must a student be able to write about such topics at a specified level of clarity and sophistication? How high is that level?

When considering such cognitive learning goals, it is important that policy makers consider the attainments of youngsters in other countries, as well as in other states. Studies that compare the performance of U.S. students with those in other industrial nations consistently rank our students near the bottom. It is thus useful to include comparisons with foreign students for two reasons. If students in other countries perform at a level superior to ours, it indicates that we, too, can do better. Another reason is that in a global economy, American employers find they must compete with the best in the world.

The definition of a "good education" should not overlook other attainments and skills we want our students to possess. We want to foster in them qualities of civility, open-mindedness, tolerance, and perseverance. Surely, we want to pass on to future generations an appreciation for the rights and responsibilities of freedom, and the difficulty and significance of its preservation.

Cognitive achievement is not the only area in which goals are needed; other important outcomes are improved student completion rates—clearly related to the education students receive—and successful transition to the job market or post-secondary education. In most cases, education goals focus on outputs and allow schools the flexibility to achieve them. Yet some important educational processes—or "inputs"—are so intrinsic to a quality education that goals to also enhance them might also be appropriate. Two examples:

- Increasing the percentage of students enrolled in advanced courses (or in the academic "track") and

- Improving the quality and performance of teachers.

Legislators alone cannot determine such goals. Other state and local officials, education professionals, parents, employers, and other citizens must be included, both to contribute their respective insights and to create a shared sense of direction.

What form will this participation take, and how will it mesh with the traditions and present structures of individual states? Responsibility for broad education goals for their children and their schools is the proper and necessary role of elected officials in every state.

As these goals are set, a strategy for achieving them must be devised. Such a strategy must be multi-dimensional, changing governance arrangements when necessary, making programmatic adjustments and providing needed resources. Key features of such a strategy are outlined in the companion handbook, *State Policies for School Restructuring*.

Assessment

Most states have some mechanism to assess educational progress. Since "progress" is measured only in relation to specific goals or targets, a new set of goals implies substantial change in existing assessment methods. A number of states use new goals to make changes that often include complex knowledge in different subject areas, writing ability, and dropout reduction.

The generally deplorable state of available commercial tests is the primary difficulty of measuring progress toward enhanced goals for cognitive learning. According to a widely cited study by testing critic John J. Cannell (1987), all 50 states reported that their students scored above the national average on standardized commercial achievement tests. One reason noted by Cannell and others is the use of outdated "norms" that compare today's students against a mythical group of "typical" students who took the test when it was created, in some cases many years ago. When compared with the antiquated norm group, today's students often benefit from "coaching" by educators familiar with the nature of the test questions. According to Cannell, today's test-takers also tend to care more about how well they do because their score often brings consequences, either for themselves or for educators—who seek to impress on them the importance of scoring well—while those who took the experimental test to establish the "norm" had little incentive to do their best. Parents and policymakers may be lulled into a false sense of comfort by results that seem to indicate that students are above the national "average."

A second problem with commercial tests, one not easily solved at the state level, is that they do not encourage comparisons between states, only to a contrived "norm." It is essentially impossible to compare the performance of students when states use different tests. Moreover, when states switch from one test to another, they lose their historical "baseline" and are unable to track improvement over time.

A third problem is that most widely used tests overemphasize simple sub-skills and isolated facts at the expense of more sophisticated analytic processes. To address this, traditional multiple-choice exams have begun to include longer, more thematic or "open-ended" questions (which offer no answer from which to choose). Several states are considering more experimental methods, intended to function as "authentic" measures of students' abilities. These are grouped into two general categories: performance tests and portfolios, both of which are discussed below. The states profiled in this handbook, South Carolina, Vermont, California, and Kentucky, either use or experiment with one or more of the newer assessment methods.

Some groups embrace such "authentic" assessments in the hope that minority test performance will improve. However, there is no evidence that students who score poorly on tests of simple skills or factual knowledge do better when assessed on writing ability or on sophisticated analytic or creative tasks. According to a RAND Corporation study, the small amount of available evidence seems to indicate the opposite (Feinberg, 1990). In a sense, the whole question of who is likely to score best is irrelevant. New cognitive measures are needed based on what is most important for students to learn.

New Cognitive Assessments

California redesigned its testing program to measure students' progress in four of the seven subject areas outlined in new curriculum frameworks: English/language

arts, mathematics, science, and history/social science. There are currently no plans to assess students in the three other subject areas: foreign language, art, and physical education. To assess the kinds of sophisticated thinking abilities envisaged in its frameworks, California added two new measures: open-ended questions that do not offer answers from which to choose and "performance tests."

Performance tests intend to assess the degree to which students apply their learning to consequential tasks, often in realistic settings. The most common performance task already appraised by many states, including California, is writing. South Carolina, for example, requires students to write an acceptable essay as part of an exam they must pass to graduate from high school. Performance tests are being developed in California in other subjects as well. These include one that requires students to conduct experiments in science and to solve extended multi-part problems in math.

Portfolio assessment is another "authentic" measure. It is drawn from the arts, where students are often judged on the basis of a self-chosen portfolio of their best work. Vermont plans to include portfolio assessment as part of a new statewide system. Yet performance tests, in fields other than writing and portfolios, are experimental, and much work is necessary before they can be counted as important components of any assessment system.

One obstacle to these newer techniques is the higher cost to administer and grade them when compared to traditional multiple-choice tests. This is due to the fact that skilled people (usually teachers), rather than computer scoring machines, must be trained and spend time grading each item. One way to address this issue is to test only a sample of the total student population rather than all students. Such a process provides data on how well the system works, and perhaps data on institutional units within the system, but it does not give parents or individual students information on how well *they* are doing.

A second issue raised by the newer assessment methods is the difficulty of maintaining grade consistency and objectivity to achieve reliable results with respect to different students. Researchers argue that this obstacle can be overcome by extensive training of assessors and by clearly designed grading criteria. They point to judges in athletic competitions as examples of a generally accepted, "non-objective" method of assessment (Wiggins, 1989).

An even greater difficulty is consistency among assessors so that student performance can be compared over time. To address this problem, the National Assessment of Education Progress (NAEP) has graders regrade a sample of the written sections from previous tests and adjust the new scores for any discrepancy. Significant differences have been noted between grades in different years (Feinberg, personal communication, October 1990). While new assessments have the potential to measure educationally important outcomes that would be difficult to measure any other way—for instance, if we wish to measure writing ability, the only way to do so is by examining actual writing samples—much can be learned from traditional multiple-choice exams. Not only are they quite good as a gauge of a student's grasp of factual knowledge, but they measure relatively sophisticated analytic and problem-solving skills.

In mathematics, the Advanced Placement, International Baccalaureate, and other exams test sophisticated higher-order thinking skills with multiple-choice questions (though some criticize them for

not measuring students' ability to apply these skills to real situations). In other subjects, such as history or literature, there is no substitute for a well-reasoned analysis buttressed by facts, in either written or oral form; both the AP and IB exams use essay questions. However, in history at least, multiple-choice tests already provide us with the necessary technology to measure important, basic knowledge in great breadth, cheaply and effectively.

Which "Levels" To Measure?

Data must be available on the performance of all education system levels that shape outcomes. These ordinarily include the state, local districts and sub-districts, schools, classrooms, and individual students. Information gathered must be distributed to the levels in a comprehensible, usable form. Since 1984, California has distributed to individual schools and districts yearly performance reports that include information on how well the schools do—compared to past performance and to their peers—on such indicators as state test scores and dropout rates. Students need reliable periodic reports on how well they are doing compared to standards that are consistent over time and uniform for all students. Students and parents should be made aware of trouble early, so they can adjust behavior while the situation is reversible.

Lack of Comparative Data

One of the most troubling gaps in our present assessment effort is a near total lack of comparable, state-level education-outcomes data. Each state's assessment effort is a unique result of the interplay of its own programs and priorities. Different state definitions and measures of dropouts, attendance, success after school completion, and other commonly used outcome measures make state-level comparisons difficult and state-to-international comparisons nearly impossible.

To address this problem, California plans to report state-level achievement data for eighth-grade math in 1991 and fourth-grade math and reading and eighth-grade math in 1993. There are no current plans or funds for state-level reporting in later years. Further, Congress prohibited even the voluntary use of NAEP tests, or data from the nationally administered test, for comparisons below the state level. In an important development, the Council of Chief State School Officers and other state groups, in conjunction with the National Center for Education Statistics (1990) developed a common definition of dropouts. Twenty-nine states now collect information under the new definition, and NCES plans to begin reporting comparable dropout data in the spring of 1993. In another example, the Southern Regional Educational Board tracks the performance of students in member states on the AP exam. State scores are reported as the percentage of students taking the exam that score at a level that qualifies for college credit (normally at least a three on the five-point AP scale).

For those who want to make comparisons at the international level, the situation is even bleaker. Intermittent comparisons have been available from private groups, including the Educational Testing Service and the Belgium-based International Association for the Evaluation of Educational Achievement. Though ETS intends to administer new international tests in 1991, and IEA plans assessments in 1994 and 1998, neither private groups nor the federal government has a reliable, long-term program for future comparisons. In a bold move, Colorado funded participation of enough of its students in the 1991 ETS international assessment that the test allows

direct comparisons between the performance of Colorado students and those in other nations.

As we seek to improve the U.S. education system, we should remember that other countries are not standing still. Without a systematic method of international comparison there is no way to ensure we are not falling further behind, much less catching up, with leading economic competitors such as Japan and Germany.

Incentives

Once clear goals and a sound way to assess progress toward them have been determined, the final step to design an accountability system is to ensure that consequences follow from good and bad results. Accountability in any organization rests on reasonable, explicit goals that are widely understood and that people are responsible to achieve. Examples of positive consequences include public praise or improved status and material benefits, such as a bonus or raise—or a combination of the two, such as a promotion. Less pleasant incentives may take the form of unfavorable public attention, job reassignment, even decreased income or dismissal. The purpose of such sanctions is not vengeance or punishment; it is to improve the likelihood that the results the organization seeks are achieved.

Public awareness of progress (or its absence) is a powerful incentive in its own right. This awareness spawns reporting techniques such as "school report cards" that provide information on individual schools and are distributed to the local news media. California, Illinois, and several other states created report cards for individual schools that include performance on state tests and other outcomes data. Though some argue that educators have an inherent interest in improved performance—and how they themselves are perceived—an increasing number of states find it useful to add more concrete incentives. "The nation's governors have called for 'real rewards' for education professionals who succeed and 'real consequences' for [those] who fail to do so" (National Governor's Association, 1989, p.2).

One option is to provide financial rewards to schools or individuals for exceptional performance. South Carolina and Florida, among others, have programs that recognize superior performance with monetary awards to schools. South Carolina also has award programs for individual teachers and school principals. Some states create "academic bankruptcy" provisions that permit intervention by state officials in cases where local performance does not reach minimum levels. These systems tend to have several stages of intervention, beginning with early warnings that result in heightened scrutiny and additional financial and technical assistance. Continued lack of improvement triggers more severe consequences. In Kentucky, state-assigned experts make high-level personnel changes at local schools if there is continued poor performance. New Jersey and Iowa are the only states to date that use such provisions to take over management of a school district. It is still too early to tell what the effects of these takeovers will be.

One limit to current "bankruptcy" sanctions is that even the most radical result in direct consequences only for upper-level administrators and school board members; they bear little consequence for rank and file employees, for children, or for parents. It should also be recognized that large, sweeping sanctions may be difficult, in political terms, to implement and they provide little day-to-day incentive.

Experimentation with finer and more precise types of sanctions is clearly one direction for future reform activity.

A crucial question for both sanctions and rewards is whether they target the school or the district. This decision cannot be made isolated from other aspects of reform. If "bottom-up" restructuring results in the school site as the primary unit of educational reform, as in Kentucky, incentives should be tailored to performance at that level. If the local district is the key unit of responsibility, it should be the unit held accountable.

Fairness

One important way to ensure fair incentives is to take into account various circumstances schools face that are beyond their control, such as students from disadvantaged backgrounds. Such adjustments must be carefully designed to prevent institutionalized poor performance among certain groups by creation of lower standards for some and higher standards for others. Incentives should be created both for improvements relative to a "peer" group of schools (or districts) with similar characteristics and for improvements compared to the entire group. South Carolina's incentive programs, described below, are viewed as fair and effective by teachers and school personnel.

Market Accountability

Some reformers argue for an entirely different kind of accountability system, one that relies on market mechanisms instead of government decisions, to provide positive and negative consequences based on performance. John Chubb of the Brookings Institution and Terry Moe of Stanford University, for example, propose that public education be reconfigured to charter new schools with only minimal state requirements (similar to those that most states now apply to private schools). These new "public" schools could compete with established district schools for students, and state funds would follow pupils to whatever school they chose to attend. Schools would thus be "accountable" to their customers—students and parents. Chubb and Moe (1990) point to current private schools, which must compete under similar circumstances and tend to provide a superior education even to disadvantaged and minority students.

Several marketplace concerns arise when considering accountability. As of this writing no state has in place such a system, and there is no way to evaluate success. In 1990, the Wisconsin legislature created a "choice" plan that includes established private schools for up to 1,000 low-income students in Milwaukee, but the results of this small-scale effort won't be available for some time. At this writing, Wisconsin's reform law is being contested in state court on technical grounds. Irrespective of the court's decision, the program seems likely to survive in some form. A statewide system akin to the proposal by Chubb and Moe was debated for the first time when a 1990 Oregon ballot initiative was defeated.

A second concern is that a market system might loosen some community bonds (while possibly strengthening others). Without self-defeating regulation, such a system might permit schools to inculcate students with eccentric versions of history or engage in separatist practices. One answer might be to evaluate all students with a common test to set minimum standards for government support. Beyond these, parents would be responsible to hold schools to more exacting standards.

Negative Effects of Incentives

As some critics point out, difficulties arise from "raising the educational stakes" by adding incentives. As the importance of performing well on these measures increases, educators may redesign their activities accordingly. This can be good or bad. Tests that measure only factual knowledge and basic skills may tempt educators to spend too much instructional time on items of minor importance. Thus, increased accountability on the part of educators cannot be separated from improved tests so that "teaching to the test" becomes a positive activity, rather than a detrimental one.

Higher stakes may also encourage educators and students to cheat; well-publicized cases of cheating have occurred in Texas, Illinois, and South Carolina. To reduce this risk, improved test security is a vital component of any reform that stresses high performance on particular tests. Secure reporting is similarly important for other measures used to guide the application of incentives, such as attendance and dropout rates. The solution to such difficulties, however, is not to give up on monitoring progress and rewarding success. California's state testing program, though it does not provide individual scores, has received praise for its secure reporting techniques. And as many as 26 states are experimenting with more "authentic" assessments to broaden the scope of material tested (Rothman, 1990).

Cost

Spending on accountability in most states typically consists of only a tiny share of total expenditures on education. The innovative and ambitious California testing program (CAP), for example, costs only about one–half of one percent of the state's education budget; even if CAP were radically expanded in line with recommendations of California's recent education summit, it would still equal only three percent of total education spending (Alexander, in press).

Without clear goals, and ways to measure progress toward them, how will reformers know if they are headed in the right direction, much less achieving significant improvements? Without fair and systematic incentives for improved results—and intervention in cases of weak results—why should these large, sluggish enterprises change comfortable, established ways of doing things? Sound investments in accountability are a necessary component of any comprehensive reform package.

No state has yet installed an assessment and accountability system that functions entirely in the manner described above, though many are experimenting with pieces of it. The sections that follow highlight particular reforms in four states. They are not intended to judge the total accountability effort of any state. Rather, they examine those reforms that seem especially promising to meet the challenge of good accountability systems.

SOUTH CAROLINA

Since passage of its wide-ranging Education Improvement Act (EIA) in 1984, South Carolina has been known as a leader among states in its search for a comprehensive approach to education reform. The EIA created a spectrum of programs and revisions in state laws aimed primarily at improvement of students' basic skills. These included boosting academic standards and inaugurating in-service training programs aimed at improving teachers' and principals' skills and professionalism.

Also included were several provisions to assess progress toward these goals and structured incentives to encourage their attainment. To pay for the new reforms, the EIA included a one-cent increase in the state sales tax, with the proceeds dedicated to education. In 1989 that tax raised over $300 million for education improvement. This section discusses the most important assessment and accountability mechanisms enacted under the EIA.

South Carolina is now in the midst of a second wave of reform as it seeks to implement "Target 2000: School Reform for the Next Decade," passed by the legislature in 1989. Target 2000 seeks to reduce dropout rates and improve student analytic skills and also includes provisions for state "deregulation" of successful districts. The new reform act is to be funded partly by the EIA tax and partly through general revenue.

Oversight Committees

To monitor implementation of the EIA, the 1984 reform act created oversight roles for two new committees. The primary oversight committee is the Select Committee on the EIA, which consists of 12 members, including legislators, the governor (or his designee), and representatives of the Department of Education. It reviews overall reform implementation and funding and makes recommendations for improvement to the legislature.

A second panel, the Business-Education Partnership for Excellence also has duties assigned by the EIA. Though technically required to report only to the legislature, its more diverse members have seen it as their mission to report to the general public (through the distribution of press releases and reports). The partnership has a standing permanent subcommittee with a full-time staff. Under the EIA, the South Carolina Department of Education is also required to provide an annual assessment of progress made under the Reform Act. The same legislation created a new unit in the South Carolina Department of Education, the Public Accountability Division. Though legally part of the Department of Education, it was conceived as a semi-independent agency without a vested interest in any individual program or unit, with a mission to monitor implementation and progress of the reform package. The Public Accountability Division issues an annual report, "What is the Penny Buying?" (referring to the one-cent sales tax used to pay for school reform).

Assessment

As part of its effort to gauge progress toward its education goals, South Carolina currently employs two different standardized testing programs. The Basic Skills Assessment Program (BSAP) is a yearly norm-referenced test unique to South Carolina. It covers reading and mathematics in grades one, two, three, six, and eight, and writing in grades six and eight. Under the Statewide Testing Program (STP), a nationally normed commercial test is also used to compare South Carolina with other states. It is not clear, however, how valid this "comparison" is, given the deceptive nature of many such "national norms."

In 1989–90, after the legislature passed an appropriation long sought by the oversight committees, the STP switched from an outdated (1981 norm) version of the Comprehensive Test of Basic Skills to the 1989 edition of the Stanford Achievement Test. The latter is a all multiple-choice test that evaluates students on language, reading, mathematics, basic science, and general social studies skills and knowledge

(South Carolina Department of Education [SC] 1986).

Incentive Programs

Exit Examinations. Beginning in the 1989–90 school year, the EIA required that all South Carolina students pass an exit examination before receiving a high school diploma. This exam is patterned after the state's Basic Skills Assessment Program; students must pass all three sections, including writing, to graduate. Students take the exam for the first time in 10th grade. If they fail any of the sections, they have one more chance in the 11th grade and two chances in the 12th to pass the parts they failed.

According to the South Carolina Department of Education, the number of 10th grade students passing all three sections of the exit exam steadily increased since it was first administered in the fall of 1986. Table I below indicates students' success rates.

Table I

Percentage of 10th Grade Students Passing All Three Sections of the Exit Exam on Their First Attempt

1986	55%
1987	58%
1988	63%
1989	66%

South Carolina Department of Education

A 1989 survey of 11th graders by an independent polling organization revealed that 68 percent of students believe that most of their fellow students felt compelled to study more because of the exit exam (SC, 1989 b).

Rewards for Schools and School Professionals. The EIA also established several incentive programs to reward outstanding teachers, principals, and schools, and it created a system of assistance, intervention, and sanctions for districts that fail to show sufficient improvement. As noted previously, programs that reward individual teachers and principals, along with associated measures to increase their quality and professionalism, are discussed in *State Policies for School Restructuring.*

Under the EIA, student achievement gains are the primary basis for distributing financial rewards under the School Incentive Reward Program. (SIRP). To qualify for an award, individual student test-score gains at a school are used to compile a School Gain Index (SGI), which must exceed the average gain for all students in that school's category. Additional financial rewards are provided for schools that also demonstrate improved student and teacher attendance rates (though no money can be awarded without a companion improvement in test scores). SIRP funds must be used to improve programs, such as the quality of instructional materials. They cannot be used to enhance employee salaries.

Each school competes in a category with other schools similar to it in three areas—family income, average years of teacher education, and (for elementary schools) percentage of students that meet or exceed school "readiness standards" on a test. The size of incentive awards is determined on a per-pupil basis, not by school. In 1989–90 the rate was $30 per student, with a total of $4.4 million disbursed to the winners. The percentage of schools that won awards that year was 26 percent. The average amount per school was $17,000. Awards also include flags and certificates (SC a).

Teachers and principals seem generally pleased with the program; 67 percent indicate it provides additional motivation to schools and 85 percent believe that goal setting and hard work win awards. Interestingly, the mostly favorable reviews come from districts with the fewest resources and, historically, the lowest levels of achievement gains (SC 1987).

District Sanctions. The EIA mandates annual performance evaluations for each school district in the state. The process, administered by the Department of Education, concentrates on "output" measures, including results on the BSAP and STP, accreditation deficiencies, dropout rates, and attendance levels. Districts that fail to reach minimum standards on at least two–thirds of these measures are declared "seriously impaired." A review committee, in conjunction with the state Board of Education, makes recommendations for improvements, which the district has six months to implement. Technical assistance and limited financial help are also provided by the state during this stage. If the district does not show satisfactory progress the state may intervene directly and remove local officials. A district is removed from "seriously impaired" status after returning to satisfactory performance levels in its annual evaluation.

Since 1984–85, nine districts have been declared "seriously impaired;" one received the designation twice for a total of 10 designations. Improved performance by the cited districts removed the designation and the need for direct state intervention. For the first time, in 1988–89 no districts received a "seriously impaired" designation SC 1986 b). The amount authorized for these evaluations and interventions is $700,000 each year, though the amount actually spent varies somewhat depending on the need for interventions.

EIA Results To Date

Five years may be considered a sufficient period of time in which to expect some preliminary results. On the whole, though with significant gaps, some improvements seem to have occurred in the education students acquire in South Carolina. The performance of South Carolina students on the national Scholastic Aptitude Test rose 35 points between 1984 and 1989, significantly greater than the one-point increase experienced nationally. The percentage of students who qualified for college credit through exams increased from 39 percent to 56 percent of those taking the test, while the number of exam-takers doubled (SC 1989 b).

There were also improvements in student performance on the STP and BSAP. The percentage of 11th graders scoring above the national norm on the STP increased from 51 percent to 56 percent of those who took the test. On the sixth grade BSAP, the percentage of students who meet state standards increased in all subjects—by 13 percent in reading, 7 percent in math, and 6 percent in writing (SC 1989 b).

Despite these gains, significant problems remain. South Carolina has no way to measure the performance of pupils on higher order thinking skills and problem solving; and EIA's emphasis on basic skills may not encourage sufficient effort in this area. The dropout rate among South Carolina students remains at about the same level as five years ago: the percentage of pupils who complete high school was 66.3 percent in 1984 and 65.9 percent in 1989. And while South Carolina students do better than ever before on the SAT, their aggregate verbal and math score is still 65 points below the national average (SC 1989 b).

Another South Carolina problem may be tied to the heightened importance it placed on test scores. As reported in a number of national newspapers, a South Carolina teacher was caught and prosecuted for cheating on the state tests she administered to her students. Allegations that such cheating is widespread in some districts in the state have been reported on the *60 Minutes* television program and in some publications (Putka 1989).
Educators in South Carolina respond that it was their efforts, and the state's new 1986 law making such activity criminal, that brought to light the activities. They also cite other steps they have taken, or plan to take, to address various aspects of test reliability, such as a new commercial achievement test (with a new "norm") to replace the previous, obsolete one; increased funding to change BSAP questions more frequently; more complex and open-ended problems; and writing assessments, all of which are more difficult to cheat (Terry Peterson, personal communication, August, 1990). Such efforts, as well as tighter test security measures, are necessary corollaries to South Carolina's increased emphasis on testing.

Target 2000: South Carolina's Second Wave of Reform

South Carolina is attempting to address some of the remaining difficulties under its 1989 reform act, Target 2000. The new act goes a long way toward filling some important gaps in the state's accountability system. As a primary goal it sets a 50-percent reduction in the
state's dropout rate by the year 2000. The new act further requires that the state's various dropout prevention programs
be evaluated for their effectiveness. Several provisions also modify different programs and accountability mechanisms to increase emphasis on students' higher order thinking skills and problem solving.

The new reform act mandates that the state Board of Education and Department of Education take into account higher order thinking skills in the adoption of new textbooks and tests, as well as in revision of the state's own tests. It further directs the state Board of Education to study the use of non-traditional (that is, "authentic") methods of to assess such higher order thinking skills.

By 1993–94, the state Board of Education, in consultation with the select committee, must also add criteria to assess higher order thinking skills and problem solving to the indices that measure the school district performance. These new criteria will be used to guide the state's incentive programs for school improvement. By 1993–94, two changes will also be made to the School Gain Index used to award incentive rewards. Relative improvements in students' performance in higher order thinking skills will be added to the criteria for those rewards. By 1991–92 the SGI must also include dropout rates. These adjustments—and those for district incentives—are intended to indicate in the clearest terms, both to educators and to students, that the state's Target 2000 goals are not exercises in rhetoric.

VERMONT

Vermont, which previously had no statewide accountability effort, developed state education goals and proceeds with plans to develop new student achievement tests and a locally oriented reporting and accountability program. Though the legislature played an important and supportive role, by approving substantial new funding when necessary, the impetus for reform has come chiefly from the state Department of Education.

Here the focus is on the new assessment methods, though we also look at new goals and several state programs that seek to foster incentives for improvement. New assessments will include "portfolios" and "best pieces" of student work, as well as more conventional tests. According to state officials, the purpose of the new assessments is to provide both information on how students are doing and to foster worthwhile classroom activities. Two new state-sponsored accountability programs are also noted below: "School Report Days," intended to provide an opportunity for schools to discuss the new results with parents and the local community;,and the "Gift of Time," which awards challenge grants to local schools to encourage creation of accountability systems.

Goals

Established in January 1990 by the Board of Education, the new goals are the result of an extended process to reach out to the state's educators and its 380,000 citizens. Five regional meetings were held and more than 20,000 individual letters sent to solicit opinions on what the goals should be. Individual meetings were held with superintendents, principals, teachers, and representatives of the PTA. Below are the goals and a few of the 41 "Measures of Success" used to assess progress toward them.

(1) Vermonters will see to it that every child becomes a competent, caring, productive, responsible citizen committed to continued learning throughout life.

- Use of mathematics and science to reason, communicate, and solve problems;

- Knowledge of world, United States, and Vermont history; geography; language; cultures; and economics;

- Knowledgeable exercise of the rights, duties, and responsibilities of citizenship.

(2) Vermonters will restructure their schools to support very high performance for all students.

- Local and state measurement of results,

- A policymaking and governance system that holds all levels accountable for performance,

- Reduction of the dropout rate to near zero.

(3) Vermont will attract, support, and develop the most effective teachers and school leaders in the nation.

- High professional standards that are continuously renewed by the profession against state, local, and national standards;

- Proof of professional competence of teachers and school leaders.

(4) Vermont parents, educators, students, and other citizens will create powerful partnerships to support teaching and learning in every community.

- Substantial increase in parents' acceptance of responsibility for, and engagement in, education of their children;

- Increased evidence of total community responsibility for the total education of each child (Vermont Department of Education, 1990).

Many of these goals, though noble, may prove difficult to evaluate. It remains to be seen how far Vermont will go to hold

educators accountable for some of the less precise measures.

Student Assessment Timeline

The new program, still under development, calls for testing all students in writing and mathematics in grades four and eight by 1991—though not at the high school level. The cost for FY 1991, the first year of full implementation, is expected to be $600,000 for testing approximately 14,000 students. The state plans to release statewide scores for the 1990-91 school year, and scores for each school in the 1991–92 school year. To oversee the process, the State Commissioner of Education has created an advisory "conscience committee," consisting of teachers, principals, and other school officials. "Stage 2" of the assessment reforms, tentatively planned for 1994, expands the subjects tested to include science, the arts, history, and social science.

Writing Assessment

The writing assessment will be conducted in three parts—a writing sample, a "portfolio" of three writings, and a "best piece." The writing sample will be produced by students under test conditions and will be evaluated by classroom teachers trained in scoring techniques. The portfolio will be selected by the students with the help of their teachers and will be graded by trained teacher. To demonstrate a broad range of ability, as well as to encourage teachers in those classes to assign more writing, students will be encouraged not to limit their choice of written assignments to those prepared for English class. The best piece will also be chosen by the student in consultation with the teacher.

The portfolio, and to an even greater degree the best piece, is intended to emphasize the importance of redrafting and revision,and to encourage students to reach their full potential. Development of all three sections seems to be progressing on schedule, with implementation slated for 1991. A draft of the guidelines for grading portfolios was completed in the spring of 1990.

Mathematics Assessment

The mathematics assessment includes an exam created by a commercial test publisher (to the state's specifications), a portfolio, and a best piece of the student's work. According to the state Board of Education, the new assessment is intended to be in the spirit of recommendations contained in the Curriculum and Evaluation Standards for School Mathematics, adopted in 1989 by the National Council of Teachers of Mathematics (Vermont Department of Education, 1990).

The exam includes problem solving and open-ended questions, as well as more traditional multiple-choice questions. It also includes questions for the purpose of national comparison. Portfolios and best piece sections are intended to showcase students' mathematical reasoning and may include extended problem solutions, highly original problem solutions, videotapes of student presentations on math topics, and such items as math problems created by the student or entries into a log or journal kept during the course of math-related work.

To judge student work, the emphasis will be on the level of demonstrated reasoning ability rather than a right or wrong answer. The portfolios and best pieces are intended to improve classroom teaching by encouraging schools to broaden the range and depth of their mathematics instruction.

A committee of seven math teachers was created by the state Department of Education and charged to consult outside experts and

develop standards to assess the portfolios and best pieces. Drafts of the new protocols are complete, and the new measures are on schedule for implementation in 1991. Development of the more conventional math test, however, probably won't be completed until 1992.

Local Accountability

A set of new programs, built on Vermont's long tradition of decentralized authority in education, has been created to assist local communities to develop their own accountability systems. Several consultants were hired by the state Department of Education to help local school boards select commercial tests and design accountability systems.

A program entitled "The Gift of Time" awards challenge grants to communities to create examples of effective use of the new school–performance data. According to the state Department of Education, only $30,000 has been appropriated for the program's first year. Also, the criteria used to make these awards have yet to be determined. For these reasons it is too early to assess any effect this program might have.

A "School Report Day" is to be held in individual schools in participating districts every May. During that day, parents, legislators, and other concerned local citizens are invited to visit schools and ask about results, including performance in the new assessment program. To help both local schools and private citizens, a commission of teachers, administrators, board of education members, and public representatives appointed by the commissioner created a pamphlet that outlines the purpose and function of the program. Though a state-sponsored program, most project funding is being raised from private grants. According to the state Department of Education, pilot tests in two communities in the 1989–1990 school year seem to have gone well, with strong turnouts reported; over 95 percent of parents, plus others from the local community, attended this year's event. Using funding from a corporate grant, the state plans to hold School Report Days in 50 of Vermont's 251 school districts in May of 1991 (Richard P. Mills, personal communication, October 1990).

Though Vermont's proposed assessment system may be considered on the "cutting edge" of such developments nationally, it is unclear how effective its accountability programs will be to encourage effective use of this new information. Reliance on private funding for School Report Days may imply a future instability caused by a lack of reliable funding. Also, it remains to be seen how well these will be attended over time as the excitement of a new program wears off, particularly in communities substantially different from the relatively small pilot communities.

The success of the "Gift of Time" grants and other state efforts to encourage development of local accountability systems is also difficult to predict. Most likely, those districts with a keen interest in education reform and accountability will benefit the most. The state's present plans do not appear to be designed to prod other districts to develop working accountability systems. Given the inherent political barriers to holding individuals accountable, results are likely to vary substantially among districts.

CALIFORNIA

California is a recognized leader in curriculum and assessment reform. The state Department of Education attempted to

use a set of model curriculum guides and the development of a complementary assessment systems levers for the improvement of California's education system. This section focuses primarily on the state's efforts to reshape its assessment tools to align them with its model curricula for grades K–12. It also looks at the state's initiative to create "school report cards" to inform parents and communities of how well their local schools are doing.

Model Curriculum Frameworks

Beginning in 1985 with its Mathematics Framework, California conducted a thorough re–examination of curricular goals and core priorities in nearly every subject. Since then, it adopted seven new "frameworks" in the areas of math, English/language arts, physical education, history/social science, visual and performing arts, foreign language, and, in late 1990, science. Not yet content, California is on the verge of a seven-year curriculum revision cycle during whichit will redesign all frameworks. The first of these, a new math framework, is scheduled for completion in 1991.

California Assessment Program

The state's primary assessment tool, the California Assessment Program, consists primarily of multiple-choice questions that test pupils' knowledge and skills in English, mathematics and, to a limited extent, science and history/social science. At a total cost of under $10 million a year, CAP's purpose is to provide information on how well the education system teaches various aspects of the curriculum at different levels—state, district, and local (Alexander, in press).

CAP employs a statistical-sampling technique that randomly tests a small percentage of students in each school and district. Individual students take only a portion of the exam so that each subject area can be tested at greater length—that is, different students are tested on different aspects of the curriculum. The result is reliable, detailed evidence about how well various components of the system educate students. There is one major drawback, however: no results are available for individual students, or their parents, to know how they're doing.

Over the past several years California has expanded the range of material tested under the CAP and is in the process of expanding it still further. Currently, the state administers CAP tests to students in grades three, six, eight, and 12. Grades three, six, and 12 test only English and math, while grade eight also tests science and history/social science. In addition to multiple-choice sections on reading comprehension and language skills, English tests in grades eight and 12—and, beginning in 1991, a new one in grade 10—assess student writing.

CAP Revisions

The California Department of Education is reformulating the English and math sections at all grade levels, changing the eighth grade history/social science section to align it more closely with the new curriculum framework, and adding history and social science to the test in grades sixth and 12. CAP is also developing assessments for chemistry and biology—field tests in these subjects were conducted in the spring of 1990. Full implementation is expected in the spring of 1991 (California Department of Education 1990a).

When they sought to implement the new curriculum frameworks, California officials discovered that the state's testing system, particularly its almost exclusive reliance on multiple-choice exams, conflicted with the spirit of the new frameworks by encouraging broad but shallow subject coverage and unnecessary fragmentation, due to the exams' emphasis on discrete sub-skills. To further coordinate its assessment program with emphasis on higher order thinking skills, such as improved analytic writing and creativity, California is developing ways of including "authentic" assessment methods in the CAP sections on math, science, and history/social science. In this ambitious effort, California is building on its successful experience with writing assessment.

CAP Writing Assessment

California first assessed writing in 1987 as part of the eighth grade CAP English test. The twelfth grade was added in 1988. Despite early protests by some that assessing writing ability would lead to arbitrary results due to inconsistent grading, California managed statewide implementation of its system in a manner generally viewed as reasonable and fair. A recent study by the University of California at Berkeley found that the new writing tests have a positive effect on classroom instruction: 78 percent of teachers said they assign more writing; 94 percent said they assign a greater variety of writing; and the percentage of students reporting they had to write 11 or more papers in a six-week period increased from 22 percent in 1987 to 33 percent in 1988 (U.S. Department of Education 1989).

One of the keys to successful implementation, according to local officials, has been inclusion of teachers in the development effort. Through the California Writing Project, a state-sponsored network of outstanding teachers, teachers played the dominant role in designing and creating the writing test. Essays are graded systematically by trained teachers at four regional scoring centers. Surprising some critics, the cost of grading each student's writing, though higher than multiple-choice tests, is only $5.

Patterned after the Writing Project, California has several other functioning teacher networks, including the California Mathematics Project, the California Literature Project, and the new California Science Project. Using these established networks, California seeks to develop similarly "authentic" assessment tools in math, science, and eventually history/social science, though there are no plans yet to evaluate literature curricula.

New Mathematics Assessment

Authentic testing in mathematics involves the addition of three new assessment methods: short, open-ended questions; "investigations"; and portfolios. Open-ended questions have appeared on the 12th grade CAP test since 1987. In 1990 results from the open-ended questions were reported for the first time as part of the school, district, and state scores. Ten percent of each student's section of the test came from the open-ended questions.

"Investigations" are 40 to 60 minute problem-solving projects that students solve individually or in small groups. These are administered by classroom teachers after proper training. They are currently in the experimental stage. Portfolio assessments of student portfolios work are still being developed.

New Science and History/Social Science Assessments

Three methods of more authentic science assessment are being tested: thematic/conceptual multiple-choice questions, open-ended questions, and performance tasks. Open-ended and thematic multiple-choice science questions, successfully piloted in 1990 on the sixth grade CAP, will be implemented statewide in 1991. Performance tasks were field tested in the spring of 1990 at the sixth grade level and will be tested in the fall at the 12th grade level. Two examples include asking students to build an electric circuit and then predict and measure its conductivity and requiring students to perform chemical tests on lake water to determine why fish are dying. Implementation of the sixth grade performance tasks is planned in 1991 and 12th grade performance tasks in 1992.

Because development is at an early stage, there are no plans to implement or field-test authentic assessment methods in history and social science. Methods under consideration include portfolios of student work; performance tasks, such as oral presentations and debates; writing assignments, both short and essay-length; and integrated assessments that may consist of multiple assessment methods and involve interdisciplinary skills.

School Report Cards

California's efforts to encourage use of CAP data to improve student performance focus on state compilation of data on the progress of individual schools. As widely reported in the media, "performance reports" for each school are compiled by the state Department of Education on performance measured by the CAP and a number of other indicators, including dropouts and student attendance. These indicators show the school's performance over time and compare it to schools with similar student populations and to all schools. The targeted audience for the performance reports is school authorities, and the format is almost indecipherable to laymen.

Proposition 98—a ballot initiative adopted in November 1988—mandates that, in exchange for increased education funding, local school boards, in cooperation with the state Board of Education, issue "school accountability report cards" for each school. The state Board of Education adopted a "model" report card in March 1989. As required by the initiative, the first cards were distributed in November 1989. By November 1990, and at regular intervals thereafter, the board must compare these with its "model" report card to see whether they conform, though no clear consequences are established if they do not.

Each report card must include, in understandable format, an introductory statement by the principal and information on how the school rates in 13 categories: student achievement, attendance and dropout levels, safety, class sizes, community involvement, facilities, and others. Because the format is designed by individual districts, readability and utility vary. Information on various "inputs," services, and processes accompany CAP scores and other outcomes. The school distributes copies to the local news media. An additional notice must be sent to the parents of each student to inform them where they may obtain free copies of the report card, or an optional shortened summary upon request.

At this writing, California was in the midst of a political battle between its governor and its elected school superintendent that has thrown the educational system (especially its accountability efforts) into a state of confusion. In August of 1990, and again in October of the same year, the governor

vetoed all funding for the CAP. Since governor-elect Pete Wilson seems committed to full funding for the CAP, negative results may be limited to a single missed CAP cycle. This highlights the importance of developing long-term consensus among the major political influences in any state. Otherwise, even innovative and promising reform may be interrupted for political reasons.

Other Accountability Mechanisms

In its fifth year of operation, California's School Recognition Program publicizes schools that achieve either outstanding or improved performance relative to schools with similar socio-economic backgrounds. Measures used are contained in each school's performance report (including CAP scores, student attendance, numbers of students enrolled in advanced courses, and so forth). Approximately 10 percent of California schools were recognized under the program in 1989-90 (California Department of Education 1990b). There is widespread state and local media coverage of award winners.

At this writing, California has no incentive programs that use concrete rewards or sanctions. In the first two years of the current wave of reform, 1984–85 and 1985–86, a financial incentive reward program did exist, but has since lost funding due to opposition from the state superintendent and lack of a strong political base. The program, widely known as "Cash for CAPs," awarded grants of $50,000 to $60,000 to successful schools (Carol Kennedy, personal communication, October 1990). It was roundly criticized by the state superintendent and the news media for its emphasis on competition, though no major abuses occurred. Schools could use the funds on anything but salaries. No alternative system was created to replace it.

Results to Date

Indicators are somewhat mixed after several years of reforms focused heavily on curriculum and assessment. Over the last four years, CAP scores remained basically flat for grades three, six, and 12, while rising slightly for grade eight. Because of a significant change in the way the state measures dropouts, there are no recent trend data to assess. SAT verbal scores have remained about the same while the average math score increased slightly. Advanced Placement scores are a bright spot, with the percentage of all seniors who qualify for college credit doubling between 1984 and 1988. In discussing California's performance, it should be noted that over the past decade the percentage of students from ethnic minorities—traditionally lower scoring populations, with the exception of Asians—has increased to the point where white students are now the minority. Because of this phenomenon, while the scores of most ethnic groups have increased somewhat, overall scores are almost unchanged.

KENTUCKY

In the spring of 1990, the Kentucky legislature was compelled by the state's highest court to reconstruct from scratch the state's entire primary-secondary education system. While courts in other states ruled their state's method of financing education produced unconstitutional inequities, Kentucky is the first in which the courts struck down the entire education system as inequitable and inefficient.

The result of the court's action was passage of a sweeping reform act that redefines and reconfigures the system, including programs, services, and funding. The new act is based on a comprehensive set of underlying assumptions about the proper functioning of an education system and the proper relationship of the state to localities and to individual schools. These assumptions include:

- The need for explicit statewide goals for improvement of educational outcomes.

- Individual schools as primary units of reform, thus vastly increasing their need for flexibility and freedom from state and district regulation.

- Local districts play an important, but secondary, role in reform implementation.

- A useful progress evaluation system should be devised, tailored to the new levels of authority—the individual school and district. It is the state's role to undertake development of such a system.

- A system of incentives with consequences for adequate and inadequate progress toward the new goals must be created, a system tailored to school and district levels.

- The state should assist schools to develop programs to achieve the new goals, but not tell them how to do so.

- Timetables should be developed to govern when different components of the new system will fall into place.

This section looks primarily at accountability provisions of the Kentucky reform act, those dealing with goal-setting, assessment of progress, and consequences for success or failure. Other parts of the act, particularly those that involve decentralization of state authority to school improvement councils and staff professionalism, are examined in the

New Bodies to Monitor Reform

To oversee successful implementation of the reform act's provisions over the next several years, the act creates an office of educational accountability that reports to the legislature. It also creates an advisory committee for educational improvement that will advise the governor, the state Board of Education and the Department of Education on issues related to accountability. The committee will be composed of citizens, parents, teachers, and education administrators.

Goal Setting

The new reform act's intent is development in all Kentucky students of a number of "capacities," including communication skills appropriate for today's complex society; the knowledge needed to make informed economic, social, and political decisions; and the skills required to compete favorably with students in other states.

It then directs the state's council on school performance standards—a panel of business people, education leaders, and private citizens—appointed by the governor to frame six goals in "measurable terms" by December 1991.
The goals are:

(1) Schools shall expect a high level of achievement of all students.

(2) Schools shall develop their students' ability to:

- Use basic communication and mathematical skills for purposes they will encounter throughout their lives;

- Apply core concepts and principles from mathematics, sciences, arts, humanities, social studies, and practical living studies to situations they will encounter throughout their lives;

- Become self sufficient individuals;

- Become responsible members of a family, work group, or community including demonstrating effectiveness in community service;

- Think and solve problems in school situations and in a variety of situations they will encounter in life; and

- Connect and integrate experiences and new knowledge from all subject matter fields with what they have previously learned, and build on past learning experiences to acquire new information through various sources.

(3) Schools shall increase their students' rate of attendance.
(4) Schools shall reduce their students' dropout and grade retention rates.
(5) Schools shall reduce physical and mental health barriers to learning.
(6) Schools shall be measured on the proportion of students who make a successful transition to work, post-secondary education, and the military.

By July 1993, the state Board of Education is required to create and disseminate to each school model curriculum frameworks based on the council's goals and recommended outcomes. Important decisions on the elaboration and implementation of these goals have been delegated to other bodies and are yet to be made. We cannot yet know how this deferral will affect the success of the act.

Assessment

By 1995–96, a deadline some have criticized as too far in the future, the state Board of Education must implement a statewide assessment program to measure the success of each school. The reform bill requires that this program include "performance-based" assessment methods. While this program is developed, an interim test will be administered to a sample of students to serve as a baseline or determine a schools' success under the new reforms. The act requires that the interim test be administered beginning in 1991–92 in grades four, eight, and 12 and that it assess students' skills in the areas of reading, math, science, and social studies. Local school boards, which are currently required to publish annual school performance reports, will include these test results.

Accountability

Using these new assessment tools, the act requires the state Board of Education to administer a system of rewards and sanctions to schools based on their level of improvement. The board must assess each school's and district's progress toward the act's goals by measuring the number of "successful" students (a term it must also define). This will be accomplished by noting a school or district's level of improvement over a two-year period.

Schools. Beginning in 1993–94, schools that show an increase in "successful" students (to be defined by the board) receive a

financial reward that they can spend as they see fit to improve education programs. Schools that show a decline in "successful" students will face one of two possible sets of consequences. If the decline was smaller than 5 percent they will be required to develop a school improvement plan, become eligible for extra funds, and be assigned one or more "distinguished educators" to provide advice and to monitor their efforts at school improvement. Distinguished educators are a new class of expert education advisors to be employed by the state Department of Education specifically for this purpose. If a school's decline was 5 percent or greater, it will be declared "in crisis" and will encounter all of the consequences listed above. In addition, the distinguished educator will have authority to recommend the transfer or dismissal of individual school staff, and students will be permitted to transfer to a successful school of their choice, even if they choose to attend a school in a different district.

Districts. Starting in July, 1991, districts that fail to meet minimum standards to be set by the state board regarding student, program, service, and operational performance shall be declared "educationally deficient" and, in cooperation with the chief state school officer, will constitute grounds for removal of the local superintendent and school board. A new local board will be appointed to choose a new superintendent. The district will not be permitted to elect a new board until the state determines that adequate performance has resumed for two successive years.

The Future in Kentucky

Perhaps the greatest difficulty that may have to be overcome surrounds implementation, particularly in some of the poor districts that were party to the original suit. Administrators and local school board members in many of them were shocked to find that not only the funding mechanism but the entire system was overturned. And some were dismayed that the new system either took away much of their authority or transferred it to new bodies, such as improvement councils at each school (Mitchell 1990).

Despite these obstacles, Kentucky has taken perhaps the most sophisticated and revolutionary approach to education reform among all statewide efforts to date. It seeks to create a fluid, interlocking system based on sound principles whose pieces are interconnected, not only with respect to accountability, but in restructuring and other reforms as well. Whether or not they choose exactly the same paths, other states that contemplate reform will find it worthwhile to take a close look at the issues Kentucky addresses.

CONCLUSION

Any type of assessment or accountability reform should be tailored to the particular education system to which it is applied. Vermont, for example, seeks to preserve aspects of a tradition of heavy local responsibility. Other reforms undertaken by the state, such as restructuring and parent enabling, must also be reflected in the accountability system. Reforms that encourage site-based management, one of the most popular contemporary approaches to restructuring, devolve a significant degree of authority to the school building level. A sound accountability system should collect information on the performance of individual schools and apply at least some of its rewards or sanctions at that level. Any effort to reform education must examine the entire system and make adjustments in all related areas and programs.

A good accountability system is necessarily dynamic. It provides information to policymakers at all levels in a manner relevant and actionable. It also provides such information to parents and the general public so that they can judge what changes are necessary. Consequences for progress, or the lack of it, encourage educators and institutions to adjust their behavior as necessary, to provide the maximum degree of fluidity and self-correction. Progress toward established goals is the result.

REFERENCES

Alexander, Francie. Forthcoming. "Accountability and Assessment California Style." In *Education Reform in the 1990s*, edited by Chester E. Finn, Jr., and Theodor Rebarber. New York: Macmillan.

California Department of Education (Calif. DOE). 1990a. *California: The State of Assessment*. Sacramento: Calif. DOE

California Department of Education. 1990b. *Award Winners Announced—California Distinguished Schools*. Sacramento, Calif. DOE. News release.

Cannell, John J. (1987). *Nationally Normed Elementary Achievement Testing in America's Public Schools: How all Fifty States are Above the National Average*. Daniels, W. V.: Friends of Education.

Chubb, John E. & Moe, Terry M. 1990. *Politics, Markets, and America's Schools*. Washington, D.C.: The Brookings Institution.

Council of Chief State School Officers (CCSSO). 1989. *Results in Education: 1989*. Washington, D.C.: CCSSO

Feinberg, Lawrence. 1990. "Multiple-choice and its Critics." *The College Board Review*, 157 (Fall) 13–31.

Mitchell, Cynthia. 1990. "Kentucky Putting School Reform to the Test." *Atlanta Journal-Constitution*. 27 March

National Center for Education Statistics (NCES) 1990. *Monthly Information Overview*. Washington, D.C.: Nov.

National Governor's Association (NGA). 1989. *State Education Indicators: 1989* Washington, D.C.: NGA

Putka, G. 1989. "Cheaters in Schools May Not be Students, but Their Teachers." *The Wall Street Journal*, 2 Nov.

Rothman, Robert. 1990. "New Tests Based on Performance Raise Questions." *Education Week* 10 (12 Sept.) 1, 10, 12.

South Carolina Department of Education (S.C. DOE) 1987. *The School Incentive Award Program: Results of the November 1987 Survey of School Staff*. Columbia, S.C.: S.C. DOE

South Carolina Department of Education (S.C. DOE) 1989a. *South Carolina School Incentive Reward Program*. Columbia, S.C.: S.C. DOE, Nov.

South Carolina Department of Education (S.C. DOE) 1989b. *What is the Penny Buying for South Carolina?* Columbia, S.C.: S.C. DOE, Dec.

U.S. Department of Education (DOE). 1989. *Learning to Write in our Nation's Schools.* Washington, D.C.: DOE

Vermont Department of Education (Vt. DOE). 1990. *Vermont's Education Goals.* Montpelier, Vt.: Vt. State Board

Vermont State Board of Education (Vt. DOE). 1990. *Mathematics, Portfolios for Instruction and Assessment.* Montpelier, Vt.: Vt. State Board, 20 Feb.

Wiggins, G. 1989. *Questions and Answers on Authentic Assessment.* Sacramento, Calif.: California Department of Education, unpublished paper.

About the Author

Theodor Rebarber is an OERI associate at the U.S. Department of Education. He was previously a research associate at the Vanderbilt University Educational Excellence Network. With Chester E. Finn, Jr., he edited *Education Reform in the 90s*, forthcoming from Macmillan Publishing.

The opinions and observations expressed in the series, *Better Education Through Informed Legislation*, are those of the author and do not reflect the policies or opinions of the U.S. Department of Education.

Forthcoming

State Policies for School Restructuring—third in the series, *Better Education Through Informed Legislation*—explores state-level policies that aim to achieve education goals by reorganizing the way learning happens. *Restructuring* discusses philosophy; considers decision-making, curriculum, instruction, and teaching reforms; and explores whether such popular policy responses as site-based management and charter schools contribute to reform goals. It will be available from the National Conference of State Legislatures in the fall of 1991.

National Conference of State Legislatures

Denver Office
1560 Broadway, Suite 700
Denver, Colorado 80202
303 /830 - 2200

Washington Office
444 North Capitol Street, N.W., Suite 500
Washington, D.C. 20001
202 /624 - 5400

Recycled Paper

BEST COPY AVAILABLE